**جامــــــــعة الشهيد حمه لخضر الــــــــوادي**

 **كلية علوم الطبيعة والحياة**

 **قسم البيولوجيا**

 **سنة ثانية ماستر س3 تخصص التنوع الحيوي وفيزيولوجيا النبات**

**محاضرة 07 الاحصاء الحيوي2**

**اختبار مربع كاي (**$X^{2}$**) :-**

* يستخدم اختبار مربع كاي للبيانات الاسمية ، فالمتغيرات يجب أن تكون مصنفة و مقاسة بمقياس إسمي ، ويستخدم لموازنة التوزيعات التكرارية بالمتغيرات ، و يسمح بمعالجة البيانات النوعية التي تكون على شكل تكرارات لمجموعات أو أصناف معينة
* يعتبر توزيع كاي تربيع من التوزيعات الإحتمالية الشائعة الاستخدام حيث توجد له تطبيقات عديدة بدرجة يمكن معها القول أنه يأتي في المرتبة الثانية للتوزيع المعتدل من حيث كثرة تطبيقاته.

كما يعد من التوزيعات المستمرة ويعتمد على التوزيع الطبيعي في حين تكون التوزيعات التكرارية غير مستمرة لذا يكون اختبار التكرارات المشاهدة مع التكرارات النظرية (المتوقعة) ذات دقة تقريبية .

 $X^{2}=∑ \frac{(0-e)^{2}}{e}$

حيث 0 قيم المشاهدات المشاهدة

حيث e قيم المشاهدات المتوقعة

ملاحظة// تكون قيم $X^{2}$ صغيرة عندما تكون قيم المشاهدات المتوقعة قريبة جداً من قيم المشاهدات المشاهدة او الواقعة وكذلك لا تكون قيم $X^{2}$ سالبة . فكلما زادت درجات الحرية كلما قل التواء التوزيع واقترب من التماثل.



يبين الشكل السابق أن توزيع  بدرجات حرية 10 يعتبر تقريبا توزيعا متماثلا، وعندما تكون درجات الحرية كبيرة بدرجة كافية فإنه يمكن تقريب توزيع  باستخدام التوزيع المعتدل. وعلى أي حال فـإن توزيع  يعتبر توزيعا مستمرا ذو قمة واحدة.

وتقارن قيم كاي سكور الجدولية والتي تستخرج على اساس درجات الحرية ومستوى المعنوية المطلوبة فاذا كانت قيمة $X^{2}$ المحسوبة اكبر او تساوي $X^{2}$ الجدولية نرفض فرضية العدم ونقبل الفرضية البديلة اي هناك فروقات معنوية .

****

**فمثلا نجد أن قيم بدرجات حرية 10 والتي تساوي المساحة على يمينها 0.05, 0.01 من المساحة الكلية هي 18.307, 23.209 على التوالي.**

**وقيم جميعها موجبة أي أنها أكبر من أو تساوي الصفر. ويرمز لــ بالرمز حيث يمثل الدليل السفلي الأول درجات الحرية بينما يمثل الدليل السفلي الثاني المساحة على يمين % القيمة.**

**فمثلا تمثل قيمة بدرجات حرية 10 والتي تساوي المساحة على يمينها 0.05 وبالنظر للجدول نجد أن :**



**استخدامات توزيع**



يستخدم توزيع في إجراء العديد من الاختبارات الإحصائية مثل:

1- **اختبار تباين المجتمع**

الاختبارات المتعلقة بتباين مجتمع ما (وذلك لاختبار المشاكل التي تتطلب اختبار تشتت مجتمع ما)، ويتم ذلك من خلال استخدام المعادلة التالية:



ويفترض في هذا الاختبار أن العينة مسحوبة من مجتمع معتدل وذلك من خلال مقارنة قيمة المحسوبة من المعادلة بالقيمة الحرجة لـ والمستخرجة من جداول .

مثال 1

إذا عرف أن تباين قوة نشاط بذور الطماطم التي تنتجها إحدى الشركات لاتزيد عن 40% ، وتستخدم الشركة الآن طريقة إنتاج جديدة يعتقد أنها ستزيد من تباين قوة نشاط البذور، سحبت عينة عشوائية من عشرة اكياس فوجد تباينها يساوي 50% .

بافتراض أن قوة مقاومة نشاط البذور تتبع التوزيع المعتدل، اختبر الفرض القائل بوجود زيادة معنوية في التباين عند مستوى معنوية 0.01 α = .

الحل

1. **وضع الفرضيات : تباين قوة نشاط البذورالمنتجة تفوق او تساوي 40 :Ho**

 **تباين قوة نشاط البذورالمنتجة أقل من 40 : H1**

2- تحديد مستوى الدلالة (  ): وهي 0.01 .



درجات الحرية = 9 ، فإن قيمة المجدولة هي 21.666

لذا فإن قاعدة القرار هي أن يتم رفض الفرضية الصفرية H0 عندما تكون



وحيث أن قيمة لاختبار تباين المجتمع يتم حسابها كالتالي :



وحيث أن قيمة المحسوبة أقل من قيمة المجدولة، فإننا بالتالي نقبل الفرضية الصفرية H0 عند مستوى دلالة 0.01 وبالتالي يمكننا القول أن بيانات العينة تدل على أن الزيادة الظاهرة في التباين ليست معنوية عند مستوى الدلالة المحدد، والشكل التالي يوضح ذلك.

11.25

21.666

**منطقة الرفض**

 1**- اختبار مربع كاي لجودة التوفيق**

ويهتم هذا النوع من الاختبارات الإحصائية باختبار ما إذا كانت مشاهدات عينة تم اختيارها من مجتمع له توزيع احتمالي معين أو نظرية معينة.

ويستخدم هذا الاختبار عندما تكون البيانات اسمية أو على شكل تكرارات ويقصد بجودة التوفيق هنا دراسة مدى تشابه تكرارات العينة والتي تسمى عادة بالتكرارات الملاحظة Observed مع التكرارات المتوقعة Expected للمتغير موضوع الدراسة فى المجتمع الأصلى.

الفروض الإحصائية:

H0 : مجموعة المشاهدات التى تم اختيارها تتبع توزيع احتمالى معين أو نظرية معينة.

H1 : مجموعة المشاهدات التى تم اختيارها لاتتفق مع هذا التوزيع أو نظرية معينة.

**إحصاء الاختبار**

$χ\_{0}^{2}=\sum\_{i=1}^{k}\frac{(O\_{i}-E\_{i})^{2}}{E\_{i}}$

اختار احد الباحثين دراسة تنوع النخيل داخل ............فوجد بها 800 نخلة موزعة كما يلي

**استخدم مستوى معنوية = 0.05 α**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| أنواع النخيل  | غرس  | دقلة نور | دقلة بيضاء  | حمراية |
| عدد المشاهدات | 200 | 350 | 150 | 100 |
| نسبة المشاهدات المتوقعة % | 25 | 45 | 15 | 15 |

**الحل:**

**الفروض الإحصائية:**

**H0 : توزيع فصيلة الدم في العينة يتفق مع التوزيع المناظر للمدينة الأخرى.**

**HA : توزيع فصيلة الدم في العينة لا يتفق مع التوزيع المناظر للمدينة الأخرى.**

**مستوى المعنوية:**

**مستوى معنوية = 0.05 α**

**احصاء الأختبار:**

****

**حيث تمثل التكرار المشاهد للنتيجة رقم i ،**

 **تمثل التكرار المتوقع المناظر للنتيجة رقم i حيث**

****

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| أنواع النخيل  | غرس  | دقلة نور | دقلة بيضاء  | حمراية |
| عدد المشاهدات | 200 | 350 | 150 | 100 |
| التكرار المتوقعة  | 200 | 360 | 120 | 120 |

**وبتطبيق معادلة كاي تربيع للحصول على قيمة كاي المحسوبة ويتم ذلك كالتالي:**

**مناطق الرفض والقبول:**

**عند استخدام جدول مربع كاي لتعيين القيمة المجدولة لـ حيث v=4-1=3 ، وبالتالي فإن**

**القرار:**

**ونلاحظ أن لذا سوف نرفض فرضية العدم وبالتالي فإن توزيع تنوع النخيل مختلف.**

**1- في دراسة فيما اذا كان هناك ارتباط بين الاصابة بالملاريا وتضخم الطحال وجدت البيانات التالية فهل هناك علاقة ارتباط بين الاصابة بالملاريا وتضخم الطحال اختبر ذلك تحت مستوى احتمال 0.01 علماً ان قيمة** $X^{2}$ **الجدولية تساوي 5.41 ؟**

|  |  |  |
| --- | --- | --- |
| ***Total*** | ***تضخم الطحال*** | ***الاصابة بالملاريا*** |
| ***-*** | ***+*** |
| ***1483*** | ***743******(936.55)*** | ***740******(546.45)*** | ***+*** |
| ***4018*** | ***2731******(2537.45)*** | ***1287******(1480.55)*** | ***\_*** |
| ***5501*** | ***3474*** | ***2027*** | ***Total*** |

**الصفتان مستقلتان عن بعضهما :H0**

**الصفتان مرتبطتان : H1**

1. $∑ \frac{(0-e)^{2}}{e}$ **=** $X^{2}$

**546.45=** $×1483$ **=**$ \frac{2027}{5501}$$e11$

**936.55=** $×1483$ **=**$ \frac{3474}{5501}$$e12$

**1480.55=** $×4018$ **=**$ \frac{2027}{5501}$$e21$

**2537.45=** $×4018$ **=**$ \frac{3474}{5501}$$e22$

$ \frac{[740-546.45]^{2}}{546.45} + \frac{[1287-1480.55]^{2}}{1480.55} + \frac{[2731-2537.45]^{2}}{2537.45}$ **=**$ \frac{[740-546.45]^{2}}{546.45}$$X^{2}$

$+ 40 + 25.30 + 14.76=148.6$ **=**$ 68.55$$X^{2}$

**بما ان قيمة** $X^{2}$ **المحسوبة اكبر من** $X^{2}$ **الجدولية نرفض فرضية العدم ونقبل الفرضية البديلة اي الصفتان مرتبطتان هناك علاقة بين الاصابة بالملاريا وتضخم الطحال .**

**2-على فرض ان احد الباحثين اراد يجد العلاقة بين الجنسين والاصابة بالسرطان فأختار عينة مؤلفة (15) فرداً (8 ذكور , 7 اناث) وقد حصل على البيانات التالية , اختبر ذلك تحت مستوى احتمال 0.05 علماً ان قيمة** $X^{2}$ **الجدولية تساوي 3.84**

|  |  |  |
| --- | --- | --- |
| ***Total*** | ***الاصابة بالسرطان***  | ***الجنس*** |
| ***مصاب***  | ***غير مصاب*** |
| ***8*** | ***2******(3.20)*** | ***6******(4.77)*** | ***ذكور*** |
| ***4018*** | ***4******(2.80)*** | ***3******(4.2)*** | ***اناث*** |
| ***5501*** | ***6*** | ***9*** | ***Total*** |

**الصفتان مستقلتان :H0**

**الصفتان مرتبطتان :H1**

$$X^{2}=∑ \frac{(0-e)^{2}}{e}$$

**4.77=** $×9$ **=**$ \frac{8}{15}$$e11$

**3.20=** $×6$ **=**$ \frac{8}{15}$$e12$

**4.2=** $×9$ **=**$ \frac{7}{15}$$e21$

**2.8=** $×6$ **=**$ \frac{7}{15}$$e22$

$ \frac{[2-3.20]^{2}}{3.20} + \frac{[3-4.2]^{2}}{4.2} + \frac{[4-2.80]^{2}}{2.80}$ **=**$ \frac{[6-4.77]^{2}}{4.77}$$X^{2}$

$ 0.45 + 0.34 + 0.51=1.62$ **=**$ 0.32$$X^{2}$

**بما ان** $X^{2}$ **المحسوبة اقل من** $X^{2}$ **الجدولية نقبل فرضية العدم ونرفض الفرضية البديلة اي الصفتان مستقلتان عن بعضهما .**